

10/502218

PCT/JP03/12518

10 Rec'd PCT/JP

20 JUL 2004

30.09.03

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 2 年 9 月 3 0 日

出 願 番 号
Application Number: 特 願 2 0 0 2 - 2 8 6 7 6 6

[ST. 10/C]: [J P 2 0 0 2 - 2 8 6 7 6 6]

出 願 人
Applicant(s): 松下電器産業株式会社

REC'D 13 NOV 2003

WIPO

PCT

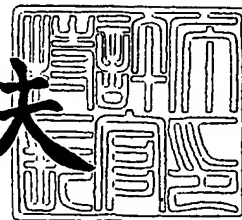
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

2 0 0 3 年 1 0 月 3 0 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 2902140123

【提出日】 平成14年 9月30日

【あて先】 特許庁長官殿

【国際特許分類】 G06K 9/32

【発明者】

【住所又は居所】 神奈川県横浜市港北区綱島東四丁目 3 番 1 号 松下通信
工業株式会社内

【氏名】 松山 好幸

【発明者】

【住所又は居所】 神奈川県横浜市港北区綱島東四丁目 3 番 1 号 松下通信
工業株式会社内

【氏名】 西澤 真人

【発明者】

【住所又は居所】 神奈川県横浜市港北区綱島東四丁目 3 番 1 号 松下通信
工業株式会社内

【氏名】 植木 千尋

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 画像補正方法

【特許請求の範囲】

【請求項 1】 2次元に配置された画素の集合体からなる画像の水平方向のヒストグラムを作成する第1のステップと、

前記ヒストグラムの値が最大になる垂直画素位置を含む文字行を検出する第2のステップと、

前記第2のステップにおいて検出された前記文字行について、前記文字行を構成する文字要素毎に、その前記画像における垂直方向の下端位置を検出する第3のステップと、

前記第3のステップで検出された前記文字要素毎の下端位置に応じて、前記画像を構成する各垂直ラインについての位置補正量を得る第4のステップと、

前記位置補正量に応じて、前記画像を構成する前記各垂直ラインを垂直方向に移動させる第5のステップとを備えたことを特徴とする画像補正方法。

【請求項 2】 前記第2のステップにおいて、前記ヒストグラムの値が最大になる垂直画素位置から、前記ヒストグラムの値が所定の範囲内となる垂直画素位置を検出することによって、前記文字行の検出を行うことを特徴とする請求項1に記載の画像補正方法。

【請求項 3】 前記第3のステップにおいて、前記文字要素毎の下端位置を直線で結んだ包絡線にもとづいて、前記位置補正量を得ることを特徴とする請求項1または請求項2に記載の画像補正方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、文書等の原稿をハンドスキャナ等で撮影した画像において生じる文字行等の蛇行を補正する画像補正方法に関する。

【0002】

【従来の技術】

従来から、文書等の原稿をスキャナ等で撮影し、撮影された文字画像を用いて

OCR (Optical Character Recognition) 処理を行って文字認識を行うための様々な技術が提案されている。

【0003】

特にハンドスキャナ等の比較的小型のスキャナを用いて原稿上をなぞって画像を撮影するという方式の装置においては、使用者が原稿上を文字の配列（以下、文字行と記す）方向に対して平行な方向に直線的になぞることは難しく、どうしても蛇行してしまい、結果的に撮影される画像が原稿と比較して蛇行したものになってしまう。このため、蛇行度合いが大きい場合、文字の歪みが影響して以後のOCR処理において文字認識率が低下するという課題があった。

【0004】

このような画像の蛇行を補正する方法としては、あらかじめ各文字画像（以下、文字要素と記す）毎に、その文字を構成する黒画素の列方向に連続したつながりの始点座標とその黒画素部分の長さからなるランレングスデータに変換し、隣接する2つの文字要素の画像を構成するランレングスデータのそれぞれの中心の列方向の位置のずれを2つの文字要素のそれぞれの中心の列方向のずれとし、そのずれに応じてランレングスデータの始点座標を変更することで蛇行した文字画像の補正を行う方法が提案されてきた（例えば、特許文献1参照。）。

【0005】

【特許文献1】

特許第3108979号公報

【0006】

【発明が解決しようとする課題】

しかしながら、前述のような画像補正方法においては、文字要素単位にシフトを行うために、文字毎の蛇行は補正できるものの、文字要素自体の歪みを補正することができず、結果的に後のOCR等の処理において、適切な文字認識ができない、という課題があった。

【0007】

本発明は、このような課題に鑑み、文字行の蛇行を補正できるとともに、文字要素自体の歪みも補正することができる画像補正方法を提供することを目的とす

る。

【0008】

【課題を解決するための手段】

本発明の画像補正方法は、2次元に配置された画素の集合体からなる画像の水平方向のヒストグラムを作成する第1のステップと、ヒストグラムの値が最大になる垂直画素位置を含む文字行を検出する第2のステップと、第2のステップにおいて検出された文字行について、文字行を構成する文字要素毎に、その画像における垂直方向の下端位置を検出する第3のステップと、第3のステップで検出された文字要素毎の下端位置に応じて、画像を構成する各垂直ラインについての位置補正量を得る第4のステップと、位置補正量に応じて、画像を構成する各垂直ラインを垂直方向に移動させる第5のステップとを備えたことを特徴としている。

【0009】

このような構成により、ヒストグラムの値が最大になる垂直画素位置を含む文字行を検出して、その行について各文字要素毎にその垂直方向の下端位置を検出する処理を行うので、画像中の全文字行について演算処理を行う場合と比較して演算部の負荷を低減することができる。

【0010】

また、このような構成により、画像を構成する全垂直ラインについて、垂直方向に移動させる補正を行うので、文字行の蛇行を補正できるのみならず、各文字要素の歪みをも補正することが可能となる。

【0011】

また、第2のステップにおいて、ヒストグラムの値が最大になる垂直画素位置から、ヒストグラムの値が所定の範囲内となる垂直画素位置を検出することによって、文字行の検出を行う構成であることにより、簡易な方法で画像中に存在している最も水平方向に長い文字行の検出が可能となる。

【0012】

さらに、第3のステップにおいて、文字要素毎の下端位置を直線で結んだ包絡線にもとづいて、位置補正量を得る構成であることにより、より簡易に文字要素

の歪みを補正することができる画像補正方法を提供することが可能となる。

【0013】

【発明の実施の形態】

以下、本発明の実施の形態について図面を用いて説明する。

【0014】

本発明の実施の形態として、本発明の画像補正方法について説明する。

【0015】

図1は、本発明の実施の形態における画像補正方法の処理ステップを示すフローチャートである。また、図2は、本発明の画像補正方法を実行するための装置の構成の一例を示す。

【0016】

図2に示したように、本発明の実施の形態における画像補正装置40は、文字や図形等の原稿を読み取って、その画像の撮影を行う画像入力部1と、画像入力部1に接続され、画像入力部1によって撮影された画像にもとづいて以降に説明する各種の処理を行うCPU2、画像入力部1によって撮影された画像の記憶を行う、フレームメモリである記憶手段3、さらに、CPU2に接続され、各種の演算された結果情報や必要な情報を使用者に対して表示する表示部5から構成されている。

【0017】

画像入力部1としては、光学的なデバイス等、公知のハンドスキャナ等に用いられるデバイスから選択したものを用いることができる。

【0018】

記憶手段3はCPU2に接続されており、その記憶媒体としては、公知のフラッシュメモリ等を用いることができる。

【0019】

表示部5は、公知のディスプレイデバイス、例えば、LCD (Liquid Crystal Display)、EL (Electro-Luminescent) およびCRT (Cathode Ray Tube) 等のデバイスから任意に選ぶことができる。

【0020】

次に、本発明の実施の形態における画像補正方法の処理ステップについて、図1に従って説明する。

【0021】

まず、画像入力部1で撮影された画像（以下、元画像と記す）が、CPU2を介して記憶手段3にロード（展開）される（ステップS29）。この元画像10の一例を図3（a）に示す。図3（a）に示した元画像10は、ハンスキャナで撮影された、部分画像を合成して得られた画像であり、ハンスキャナを使用者が手動でスキャンした際のスキャン方向と、文字の行方向とのなす角度の差が大きく影響して、文字行の方向が傾斜してしまっている。さらに、ハンスキャナを用いて原稿上をなぞった際の軌跡が、文字行方向に対して蛇行したために、図3（a）に示した元画像10は蛇行してしまっている。なお、本実施の形態において、元画像10は、2次元に配置された画素の集合体であり、それぞれの画素の輝度値は多値（256階調）の白黒画像である。

【0022】

また、本実施の形態においては、画像入力部1の撮影手段として、 256×16 画素のCCDを用い、画像を展開するための、記憶手段3としては、図3（a）において紙面に対して水平方向（横） \times 紙面に対して垂直方向（縦） $= 1000 \times 400$ 画素のフレームメモリを用いた。

【0023】

次に、CPU2は、記憶手段3に記憶された元画像10に対して、全体の画像の傾きを補正する処理を実行する（ステップS30）。本発明は、この傾きを補正する処理についてなんら限定するものではなく、公知の方法を用いることができる。例えば、特開平1-156887号公報には、複数の角度に元画像10を回転させ、行方向のヒストグラムを算出して、そのヒストグラムの幅が最も小さくなるような角度を、元画像10を回転させるべき角度であるとして決定し、元画像10を回転させる方法が開示されており、この方法を用いることもできるし、他の公知のいかなる方法を用いてもよい。

【0024】

図 3 (b) にこのような傾きを補正する処理を行った、傾き補正された画像 1 1 の一例を示す。傾き補正された画像 1 1 は、濃度値が「1」の黒画素と、濃度値が「0」の白画素とから構成された、いわゆる二値化された画像である例を示す。傾き補正された画像 1 1 は、4 つの文字行 (図 3 (b) 中の A ~ D) から構成されている。なお、本明細書中において、文字行とは、文字が記載された方向 (例えば、図 3 (b) 中においては X 軸方向) の文字要素のつながりをいう。

【0 0 2 5】

傾き補正された画像 1 1 と元画像 1 0 とを比較すると、画像全体の傾きは補正されているものの、前述した蛇行は残存していることが分かる。以下、本実施の形態における画像の蛇行を補正する方法について、説明する。

【0 0 2 6】

再び図 1 に戻って、CPU 2 は、最も行方向 (図 3 (b) における X 軸方向) に長い文字行を検出するために、記憶手段 3 に記憶された傾き補正された画像 1 1 について、図 3 (b) 中の各垂直画素位置について、各水平ライン毎に水平方向の黒画素の数をカウントして (濃度値を加算して)、積算ヒストグラムを算出する (ステップ S 3 1)。

【0 0 2 7】

なお、本明細書中においては、元画像 1 0 を構成する画像データにおける水平方向の画素の連なりを水平ラインと記し、垂直方向の画素の連なりを垂直ラインと記す。

【0 0 2 8】

図 4 は、図 3 (b) に示した傾き補正された画像 1 1 について、水平方向の積算ヒストグラムを算出した結果である。図 4 においては、横軸に画像の垂直方向の画素の位置を示し (図 3 (b) においては、紙面に向かって左上の隅を原点 O とした例を示す)、縦軸には、黒画素の数をプロットしたものである。

【0 0 2 9】

図 4 から明らかなように、本実施の形態においては、算出された黒画素のヒストグラムは、4 つの山部 (A ~ D) に分かれていることが分かる。これらの山部は、それぞれ、図 3 (b) に示した傾き補正された画像 1 1 の、前述した 4 つの

文字行（A～D）に対応している。

【0030】

続いてCPU 2は、前述のステップで算出された水平方向の積算ヒストグラムから、黒画素数が最大値となる垂直方向の画素位置を算出し、最大値を含む山部を、長さが最長の文字行であるとみなす（ステップS 3 2）。本実施の形態においては、山部Bが、黒画素数が最大値となる垂直方向の画素位置を含むので、文字行Bが最長の文字行であるとみなされる。

【0031】

さらにCPU 2は、長さが最大とみなされた山部Bについて、対応する文字行の垂直方向の幅を決定する（ステップS 3 3）。具体的には、黒画素数が最大値に対して所定の割合、R %になる垂直方向の画素位置範囲（図4において、Wで示される範囲）を最長文字行範囲Wとして決定する。Rは、隣接する文字行と分離できる値に設定する。実用的にはR = 3 0程度に設定することで確実に文字行の分離を行うことが可能である。

【0032】

ここまでの処理ステップによって、蛇行補正すべき量を決定するための垂直方向の最長文字行範囲Wを決定することができる。

【0033】

次に、最長文字行範囲Wとして決定された垂直方向の画素位置範囲において、CPU 2は、図3（b）に示した傾き補正された画像1 1上を、原点からX軸方向に順次垂直1ライン毎に、図3（b）における紙面に向かって下側からスキャンして、黒画素が初めに検出された位置を黒画素下端位置として決定する。この処理を全垂直ラインについて行う（ステップS 3 4）。

【0034】

図5は、図3（b）に示した傾き補正された画像1 1の全垂直ラインについて、黒画素下端位置を検出した結果を示す図である。説明を簡単にするために、図5には、傾き補正された画像1 1を重畳表示している。

【0035】

図5を見て分かるように、文字行B（厳密に言えば最長文字行範囲W）につい

て、文字行を構成する各文字要素毎に、最も下端の黒画素下端位置が検出されていることが分かる。

【0036】

図1に戻って、次にCPU2は、図5に示した算出結果のうち、黒画素下端位置の紙面に向かって最も下側の点を結ぶ包絡線を算出して、垂直位置ずれ量として決定する（ステップS35）。具体的には黒画素下端位置が「0」となる垂直ラインで挟まれた領域を1文字要素であると仮定して、各文字要素を構成する領域毎に、最も下側の黒画素下端位置を検出し、隣接する各文字毎の黒画素下端位置同士を直線で結んで、各垂直ライン毎に、垂直方向に補正すべき垂直位置ずれ量を算出する。図6に垂直位置ずれ量の検出結果の一例を示す。

【0037】

次に、CPU2は、図6に示した垂直位置ずれ量にもとづいて、各垂直ライン毎に垂直方向にずらすべき量を算出して、記憶手段3に記憶された各垂直ライン毎に、垂直方向にずらして処理を終了する（ステップS36）。なお、このステップの後に、表示部5が、蛇行補正された画像を使用者に対して表示する構成であってもよい。

【0038】

このように、本実施の形態の画像補正方法によれば、1文字とみなされる領域毎に黒画素下端位置を検出して、隣接する文字要素毎の黒画素下端位置同士をつなぐような包絡線を作成して、この垂直位置ずれ量にもとづいて全垂直ラインについてずらす補正を行うので、文字要素自体の歪みも改善することができる。

【0039】

図7に、本実施の形態に示した画像補正方法にもとづいて処理を行った、蛇行補正された画像13を示す。図7を見て分かるように、図3（b）に示した傾き補正された画像11と比較して、その文字行の蛇行の度合いが大幅に軽減している。本実施の形態に示したように、最も長い文字行にもとづいて補正すべき量を決定して、全画像を構成する垂直ラインについて補正を行うことにより、画像全体の蛇行を補正できることが確認できた。

【0040】

このように、本実施の形態の画像補正方法によれば、最も長い文字行を、水平方向のヒストグラム算出によって検出して、その文字行について補正すべきずれ量を検出して、全画像について補正を行うので、画像を構成する全ての文字行についてずれ補正をする場合と比較して、処理を迅速に行うことが可能となる。

【0041】

本実施の形態の画像補正方法を用いることにより、原稿の文字認識精度が向上したことを以下に記述する。

【0042】

正読率の算出は、記憶手段3に記憶された画像にもとづいて、CPU2でOCR処理を行い、その結果正しく認識された文字の割合を算出することにより行う。サンプルとしては、ランダムに名刺20枚を用いて認識を行った。撮影し、OCRに供した文字数は、電話番号390文字、メールアドレスおよびURLは1026文字である。

【0043】

まず、名刺中の電話番号の場合、補正を行わない場合の正読率に対して、本実施の形態の画像補正方法を行うことにより、約20%正読率を向上することができた。

【0044】

また、名刺中のメールアドレスおよびURLについても、補正を行わない場合の正読率に対して、本実施の形態の画像補正方法を行うことにより、約25%正読率が向上しており、より高い正読率を得ることができた。

【0045】

また、電話番号、メールアドレス、URL全体においても、補正を行わない場合の正読率に対して、本発明の画像補正を行った場合には正読率を約23%向上することができた。これは、本発明の画像補正方法によれば、文字行の蛇行を補正できるとともに、文字自体の画像の歪みも補正することができるためであると考えられる。

【0046】

【発明の効果】

以上述べたように、本発明の画像補正方法を用いれば、文字行の蛇行を補正できるとともに、文字画像自体の歪みも補正することができる。

【図面の簡単な説明】

【図 1】

本発明の実施の形態における画像補正方法の処理ステップを示すフローチャート

【図 2】

本発明の実施の形態における画像補正装置の構成の一例を示すブロック図

【図 3】

(a) は本発明の実施の形態における元画像の一例を示す図

(b) は本発明の実施の形態における傾き補正された画像の一例を示す図

【図 4】

本発明の実施の形態における水平方向積算ヒストグラムの一例を示す図

【図 5】

本発明の実施の形態における黒画素下端位置の検出結果の一例を示す図

【図 6】

本発明の実施の形態における垂直位置ずれ量の検出結果の一例を示す図

【図 7】

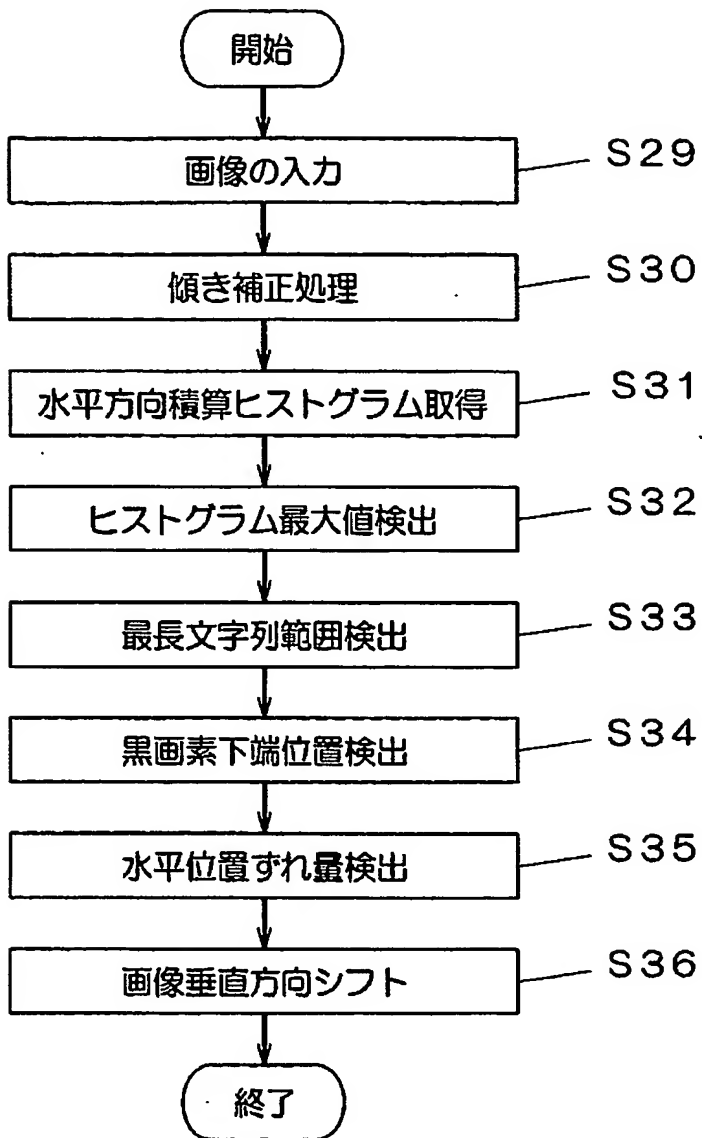
本発明の実施の形態における蛇行補正された画像を示す図

【符号の説明】

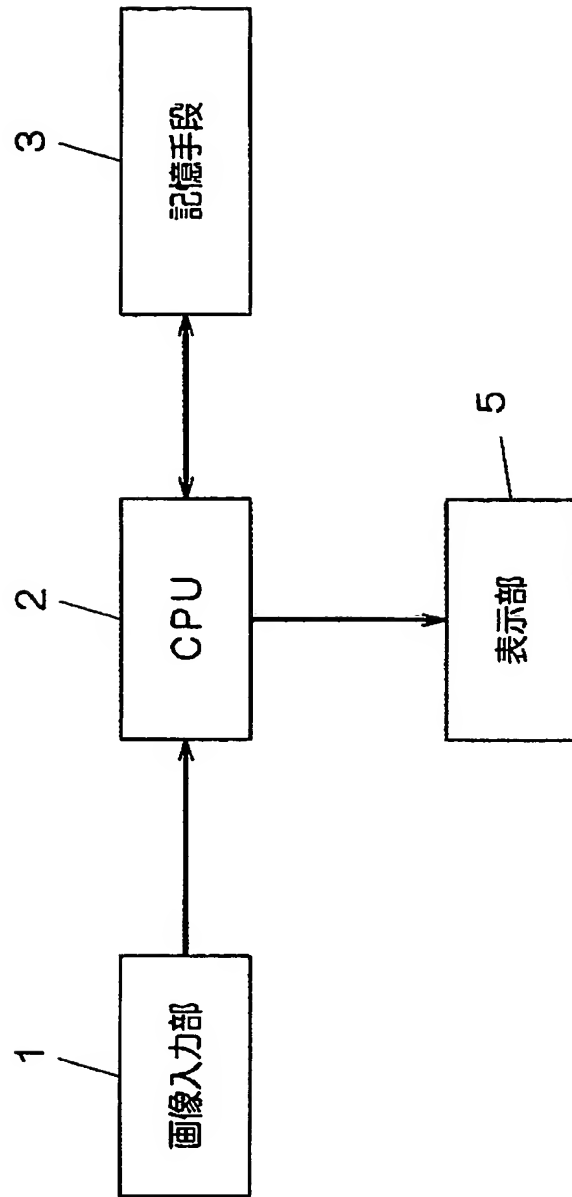
- 1 画像入力部
- 2 CPU (演算部)
- 3 記憶手段
- 5 表示部
- 10 元画像
- 11 傾き補正された画像
- 13 蛇行補正された画像
- 40 画像補正装置

【書類名】 図面

【図 1】



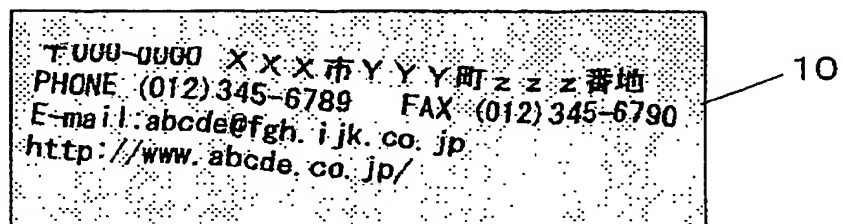
【図 2】



40 画像補正装置

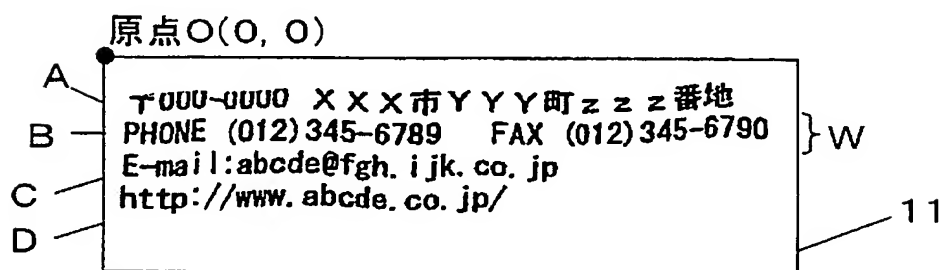
【図3】

(a)



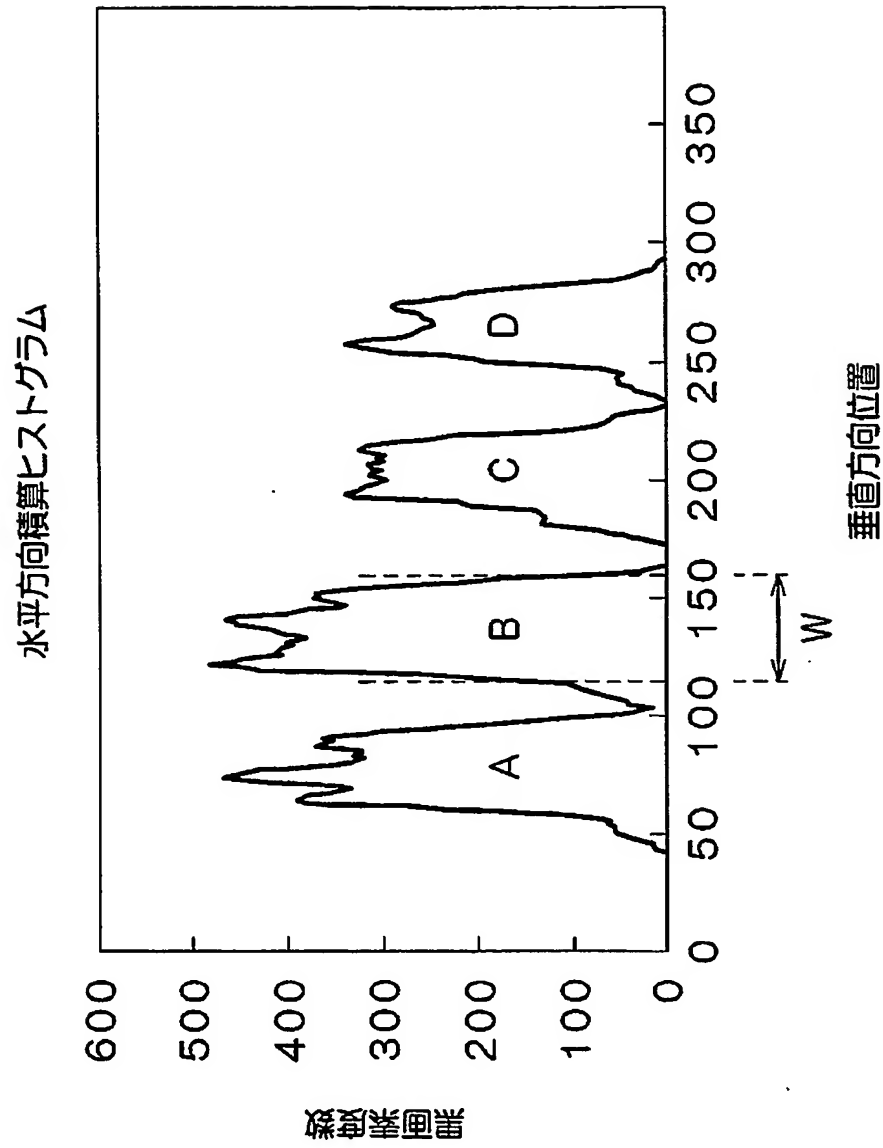
X軸方向(処理方向)

(b)

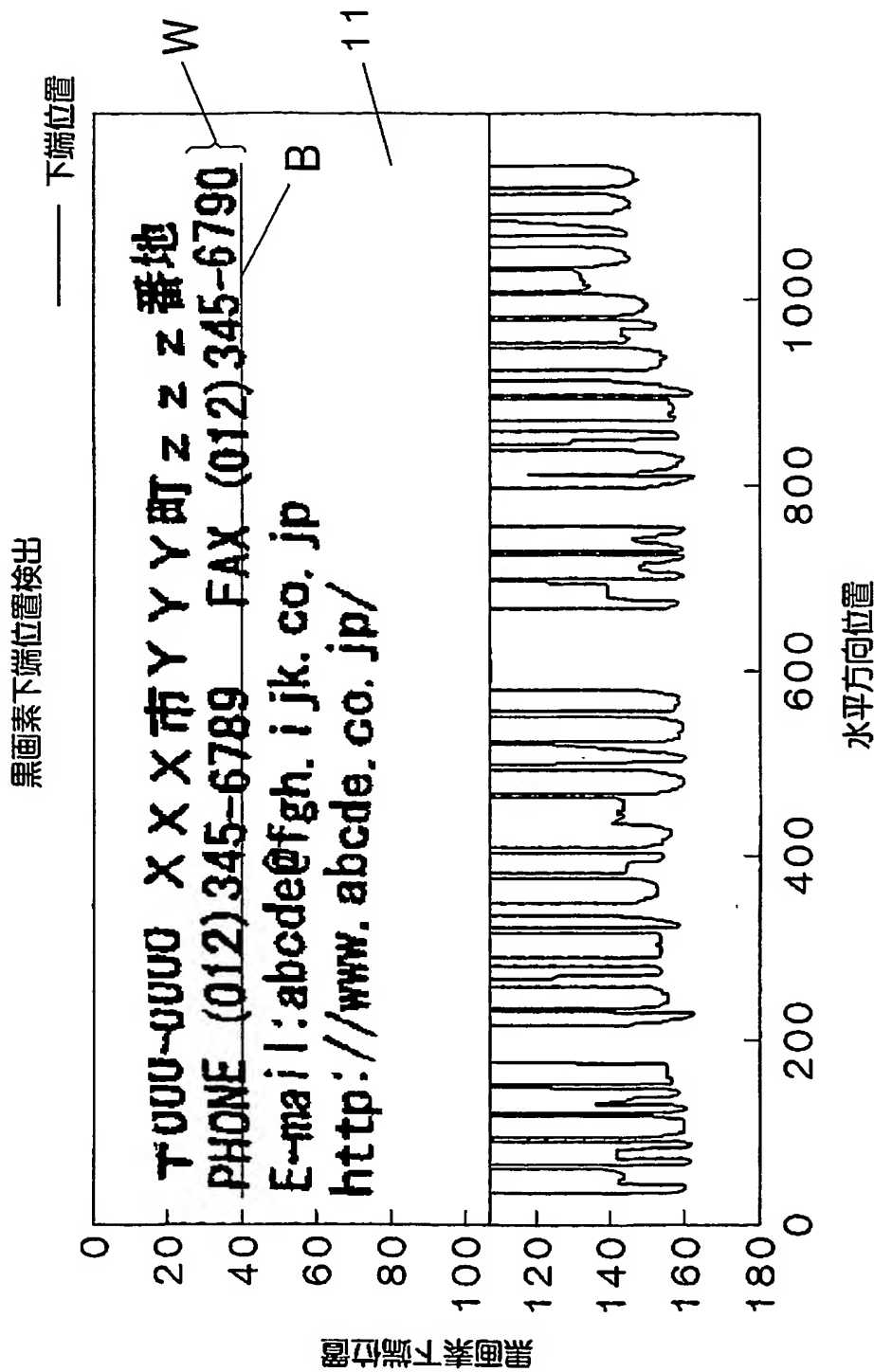


X軸方向(処理方向)

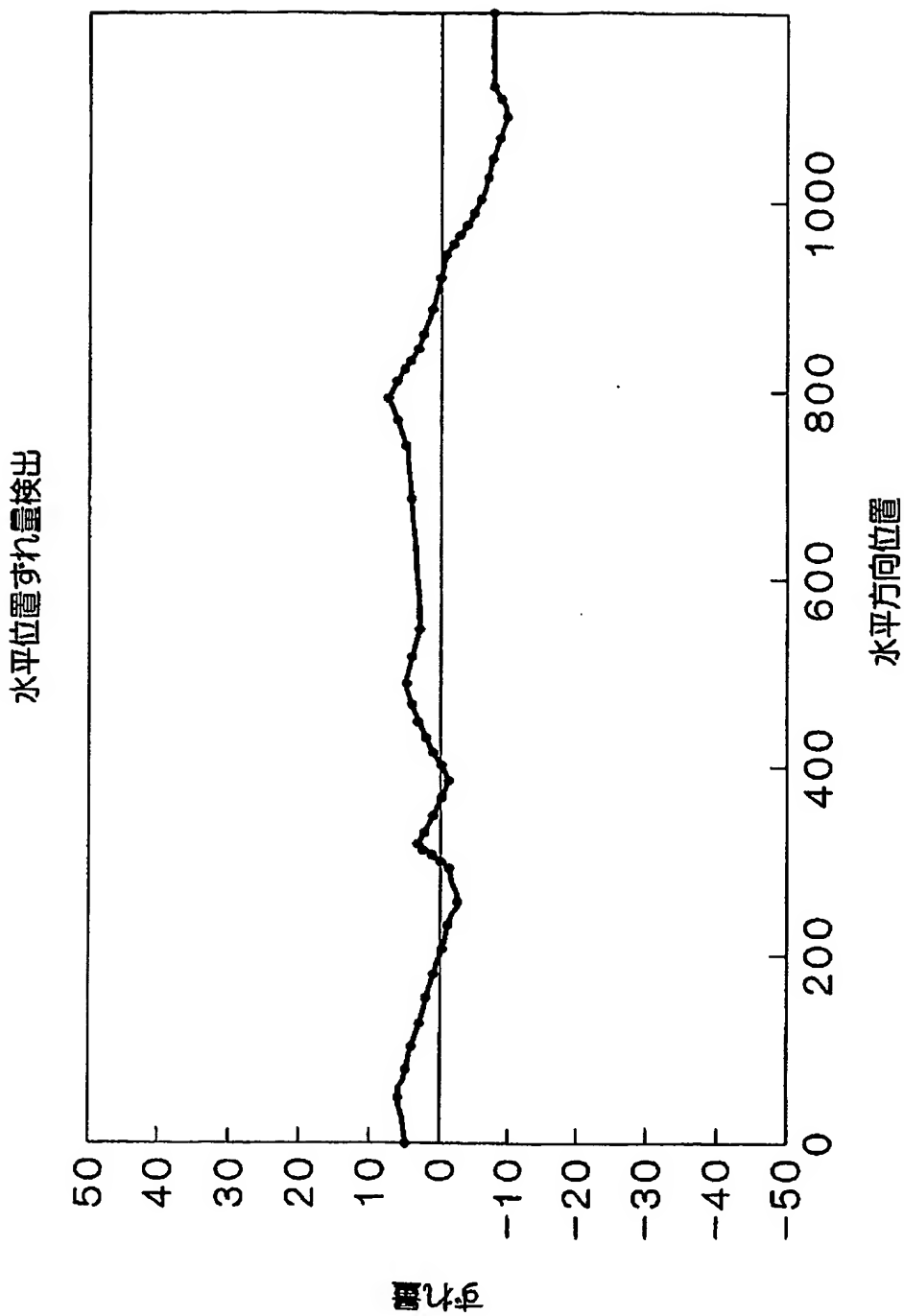
【図 4】



【図 5】



【図 6】



【図 7】

〒000-0000 X X X 市 Y Y Y 町 z z z 番地
PHONE (012)345-6789 FAX (012)345-6790
E-mail:abcde@fgh.ijk.co.jp
<http://www.abcde.co.jp/>

13

【書類名】 要約書

【要約】

【課題】 文字行の蛇行を補正できるとともに、文字要素自体の歪みも補正することができる画像補正方法を提供する。

【解決手段】 画像の水平方向のヒストグラムを作成する第1のステップと、ヒストグラムの値が最大になる垂直画素位置を含む文字行を検出する第2のステップと、第2のステップにおいて検出された文字行において、文字行を構成する文字要素毎にその垂直方向の下端位置を検出する第3のステップと、第3のステップで算出された、文字要素毎の下端位置にもとづいて、画像を構成する各垂直ラインについて、補正すべき量を算出する第4のステップと、各垂直ラインを補正すべき量にもとづいて垂直方向に移動させる第5のステップとを行う。

【選択図】 図1

特願 2 0 0 2 - 2 8 6 7 6 6

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 8 2 1]

1 . 変 更 年 月 日

1 9 9 0 年 8 月 2 8 日

[変 更 理 由]

新 規 登 録

住 所

大 阪 府 門 真 市 大 字 門 真 1 0 0 6 番 地

氏 名

松 下 電 器 産 業 株 式 会 社